

## Using T.O.M to Estimate Parameter of distributions that have not Single Exponential Family

Jubran Abdul Ameer Khuttar Al-Labban

Department of Statistics and Informations College of Computer Science and Mathematics University of Al-Qadisiya IRAQ

**Abstract:** We discuss the T.O.M (Term Omission Method) to estimate the distribution parameters that have not single exponential family of exponential family with one single distribution parameter, and compare it with different methods using Mean Square Method (MSE).

**Keywords** - Exponential families, Exponential form,

### I. INTRODUCTION:

Estimators that based on T.O.M deals with many distributions with discrete or continuous random variable that have not belongs to exponential family or it's taking exponential form. Many distributions not belongs to exponential family like, Uniform, Hypergeometric, t-distribution, F distribution, Cauchy, shifted exponential, Pareto.<sup>[1\*]</sup>

### II. EXPONENTIAL FORM:

We note that there are many definitions of such type of that representation of exponential families. Recall definitions of exponential families.

### III. DEFINITIONS:

**Def. (1):** A family of continuous (discrete) random variables is called an exponential family if the probability density functions (probability mass functions) can be expressed in the form

$$f_X(x/\theta) = h(y)c(\theta) \exp\left(\sum_{i=1}^k \theta_i t_i(x)\right), \quad x = 0,1,2,\dots \quad \dots (1)$$

for  $x$  in the common domain of the  $f_X(x/\theta)$ ,  $\theta \in \mathbb{R}^k$ .

Obviously  $h$  and  $c$  are non-negative functions. The  $t_i(x)$  are real-valued functions of the observations.<sup>[2\*]</sup>

**Def. (2):** Let  $\mathcal{G}$  be an interval on the real line. Let  $\{f(x;\theta); \theta \in \mathcal{G}\}$  be a family of pdf's (or pmf's). We assume that the set  $\{\underline{x} : f(\underline{x};\theta) > 0\}$  is independent of  $\theta$ , where  $\underline{x} = (x_1, x_2, \dots, x_n)$ . We say that the family  $\{f(x;\theta); \theta \in \mathcal{G}\}$  is a one-parameter exponential family if there exist real-valued functions  $Q(\theta)$  and  $D(\theta)$  on  $\mathcal{G}$  and Borel-measurable functions  $T(\underline{X})$  and  $S(\underline{X})$  on  $\mathbb{R}^n$  such that

$$f(\underline{x};\theta) = \exp(Q(\theta)T(\underline{x}) + D(\theta) + S(\underline{x})) \quad \dots (2)$$

if we write  $f(\underline{x};\theta)$  as

$$f(\underline{x};\eta) = h(\underline{x})c(\eta) \exp(\eta T(\underline{x})) \quad \dots (3)$$

where  $h(\underline{x}) = \exp(S(\underline{x}))$ ,  $\eta = Q(\theta)$ , and  $c(\eta) = \exp(D(Q^{-1}(\eta)))$ , then we call this the exponential family in canonical form for a natural parameter  $\eta$ .

**Def. (3):** Let  $\underline{\mathcal{G}} \subseteq \mathbb{R}^k$  be a  $k$ -dimensional interval. Let  $\{f(x; \underline{\theta}); \underline{\theta} \in \underline{\mathcal{G}}\}$  be a family of pdf's (or pmf's). We assume that the set  $\{\underline{x} : f(\underline{x};\underline{\theta}) > 0\}$  is independent of  $\underline{\theta}$ , where  $\underline{x} = (x_1, x_2, \dots, x_n)$ . We say that the family  $\{f(x; \underline{\theta}); \underline{\theta} \in \underline{\mathcal{G}}\}$  is a  $k$ -parameter exponential family if there exist real-valued functions  $Q_1(\underline{\theta}), \dots, Q_k(\underline{\theta})$  and  $D(\underline{\theta})$  on  $\underline{\mathcal{G}}$  and Borel-measurable functions  $T_1(\underline{X}), \dots, T_k(\underline{X})$  and  $S(\underline{X})$  on  $\mathbb{R}^n$  such that:<sup>[3\*]</sup>

$$f(\underline{x}; \underline{\theta}) = \exp\left(\sum_{i=1}^k Q_i(\underline{\theta})T_i(\underline{x}) + D(\underline{\theta}) + S(\underline{x})\right) \dots (4)$$

**Def. (4):** Exponential family is a class of distributions that all share the following form:

$$P(y/\eta) = h(y) \exp\{\eta^T T(y) - A(\eta)\} \quad \dots (5)$$

- \*  $\eta$  is the natural parameter. For a given distribution  $\eta$  specifies all the parameters needed for that distribution.
- \*  $T(y)$  is the sufficient statistic of the data (in many cases  $T(y) = y$ , in which case the distribution is said to be in canonical form and  $\eta$  is referred to as the canonical parameter).
- \*  $A(\eta)$  is the log-partition function which ensures that  $p(y/\eta)$  remains a probability distribution.
- \*  $h(y)$  is the non-negative base measure (in many cases it is equal to 1).

Note that since  $\eta$  contains all the parameters needed for a particular distribution in its original form, we can express it with respect to the mean parameter  $\theta$ :<sup>[1]\*</sup>

$$P(y/\theta) = h(y) \exp\{\eta(\theta)T(y) - A(\eta(\theta))\} \quad \dots (6)$$

**Def. (5):** (Regular Exponential Family): Consider a one-parameter family  $\{f(x;\theta) : \theta \in \Omega\}$  of probability density functions, where  $\Omega$  is the interval set  $\Omega = \{\theta : \gamma < \theta < \delta\}$ , where  $\gamma$  and  $\delta$  are known constants, and where<sup>[4]\*</sup>

$$f(x;\theta) = \begin{cases} e^{p(\theta)k(x)+s(x)+q(\theta)} & a < x < b \\ 0 & \text{o.w} \end{cases} \quad (7)$$

The form (7) is said to be a member of the exponential class of probability density functions of the continuous type, if the following conditions satisfy:

- 1) Neither (a) nor (b) depends upon  $\theta$ .
  - 2)  $p(\theta)$  is a nontrivial continuous function of  $\theta$ .
  - 3) Each of  $k'(x) \neq 0$  and  $s(x)$  is a continuous function of  $x$ .
- and the following conditions with discrete random variable  $X_i$ :
- 1) The set  $\{x : x = a_1, a_2, \dots\}$  does not depend upon  $\theta$ .
  - 2)  $p(\theta)$  is a nontrivial continuous function of  $\theta$ .
  - 3)  $k(x)$  is a nontrivial function of  $x$ .

Therefore, the Exponential Form is the Exponential Class without satisfying any of the conditions in above definitions or can not written in Exponential Class for discrete or continuous random variables.

#### IV. EXAMPLES OF DISTRIBUTIONS THAT NOT BELONGS TO EXPONENTIAL FAMILY:

We will take here some distributions as an example; to show how can write the p.d.f as an exponential form.

**Example 1:** Let  $X$  be a continuous random variable of Uniform distribution with parameter  $\theta$ , with p.d.f

$$f(x;\theta) = \frac{1}{\theta} \quad 0 \leq x \leq \theta$$

here we can not represent this distribution as an exponential class because  $x$  depends upon  $\theta$ , so we can rewrite the p.d.f of this distribution represented by exponential form which can be the same form of exponential class as follows:

$$f(x;\theta) = \exp[-\ln(\theta)]$$

with  $k(x) = 0$  ,  $p(\theta) = 0$  ,  $s(x) = 0$  ,  $q(\theta) = -\ln(\theta)$

**Example 2:** Let  $X$  be a continuous random variable have the shifted exponential distribution with p.d.f

$$f(x;\theta) = \frac{1}{\theta} e^{1-\frac{x}{\theta}} \quad 0 < \theta < x < \infty$$

with scale parameter  $\theta$ , and also we can here represented the p.d.f as an exponential form because random variable depends upon the parameter  $\theta$ ,

$$f(x;\theta) = \exp\left[-\ln(\theta) + 1 - \frac{x}{\theta}\right]$$

where  $k(x) = -x$ ,  $p(\theta) = \frac{1}{\theta}$ ,  $s(x) = 1$  ,  $q(\theta) = -\ln(\theta)$

note that the random variable in previous two examples (1,2) depends on  $\theta$  and the condition (1) of def. (5) not satisfying.

**Example 3:** Let X be a continuous random variable of Cauchy distribution with parameter  $\alpha, \beta$  with p.d.f

$$f(x; \alpha, \beta) = \frac{1}{\pi} \left[ \frac{\beta}{(x - \alpha)^2 + \beta^2} \right] \quad -\infty < x < \infty$$

where  $\alpha$  is the location parameter, and  $\beta$  is the scale parameter, the one parameter Cauchy distribution can be written when  $\alpha=0$  as follows:

$$f(x; \alpha, \beta) = \frac{1}{\pi} \left[ \frac{\beta}{x^2 + \beta^2} \right] \quad -\infty < x < \infty$$

The reason of why we can not represented the Cauchy distribution as an exponential families is that we can not determine the term  $k(x)p(\theta)$  or any other terms of exponential families, or some other reasons<sup>[5\*]</sup>, so we will represent the Cauchy distribution in the exponential form

$$f(x; \alpha, \beta) = \exp \left[ -\ln(\pi) + \ln(\beta) - \ln(x^2 + \beta^2) \right]$$

where  $k(x)p(\beta) = -\ln(x^2 + \beta^2)$ ,  $s(x) = -\ln(\pi)$ ,  $q(\beta) = \ln(\beta)$

### V. T.O.M WITH EXPONENTIAL FORM:

T.O.M can used to estimate the value of parameter in many cases, below we will derivative the T.O.M of distributions (with one parameter) have not exponential family or have exponential form with two cases:

Case 1: The distribution can be written as exponential class but not belongs to exponential family because unsatisfying one or more conditions of exponential families, like Uniform, student's t, etc.

Case 2: The distribution can be not written as exponential class, like Shifted exponential, Cauchy, Pareto, etc..

We will discuss each case using T.O.M as follows:

#### 5.1 T.O.M WITH CASE 1:

There are many distributions can be written as exponential class but it's not belong to exponential family. Here we will point to some reasons why such distributions not belongs to exponential family ( $p(\theta)=0$ , random variable depend on parameter, etc.) as shown in examples (1, 2), so for our case we can use T.O.M to estimate parameter of that distributions.

For sample with size n having the p.d.f  $f(x_i; \theta)$ , and for any value  $x_i, i=1, \dots, n$

$$x_i \quad y_i = e^{k(x_i)p(\theta) + q(\theta) + s(x_i)}$$

by taking the natural logarithm to  $f(x_i; \theta)$ , since  $p(\theta)=0$  we have

$$x_i \quad q(\theta) + s(x_i)$$

and by subtract  $s(x_i)$  from the final amount we have  $q(\theta)$  which is the function of  $\theta$ .

Therefore, we can define the  $q^i(\theta)$  as follows:

$$q^i(\theta) = \ln y_i - s(x_i) \quad \dots(8)$$

where  $q^i(\theta)$  represent to values that we have from previous steps of T.O.M,  $\forall i=1,2, \dots, n$ , and from eq. (8) we can educe values of  $\theta^i$  from  $q^i(\theta)$ .

The estimation of  $\theta$  can found now using the least square error from the following equation:

$$\hat{\theta} = \text{Min} \left( \sum_{m=1}^n [f(x_m, \theta^i) - y_m] \right) \quad i = 1, 2, \dots, n$$

where  $f(x_m, \theta^i)$  is the value of function  $f(x)$  on  $\theta^i$ , and  $y(x_m)$  is the observed value.

#### 5.2 T.O.M WITH CASE 2:

There are many distributions can not be written as exponential class for some reasons. Here we will point to some reasons why such distributions not belong to exponential family. Often distributions that given by the location-scale distribution families; like Cauchy distribution<sup>[1\*]</sup>, shifted exponential density<sup>[6\*]</sup>; such as the p.d.f in example 2. In this case we can call T.O.M for exponential class<sup>[7\*]</sup>.

So in our cases we can use T.O.M to estimate parameter of those distributions.

**Example 4:** In example (1) of Uniform distribution we can obtain the following exponential form:

$$f(x; \theta) = \exp[-\ln(\theta)]$$

where  $k(x) = 0$  ,  $p(\theta) = 0$  ,  $s(x) = 0$  ,  $q(\theta) = -\ln(\theta)$

since  $p(\theta) \neq 0$ , then we can use eq. (8), hence

$$q^i(\theta) = \ln y_i$$

and

$$\theta^i = \frac{1}{y_i}$$

**Example 5:** In example (2) of p.d.f we can obtain the following exponential form:

$$f(x; \theta) = \exp\left[-\ln(\theta) + 1 - \frac{x}{\theta}\right]$$

where  $k(x) = -x$ ,  $p(\theta) = \frac{1}{\theta}$ ,  $s(x) = 1$  ,  $q(\theta) = -\ln(\theta)$

since  $p(\theta) \neq 0$ , then we can use the same way of T.O.M<sup>[7\*]</sup>, thus

$$p^i(\theta) = \frac{\ln(y_{i+1}) - \ln(y_i)}{x_i - x_{i+1}}$$

and

$$\theta^i = \frac{x_i - x_{i+1}}{\ln(y_{i+1}) - \ln(y_i)}$$

**Example 6:** in example (3) the p.d.f can obtain by the following exponential form:

$$f(x; \alpha, \beta) = \exp\left[-\ln(\pi) + \ln(\beta) - \ln(x^2 + \beta^2)\right] \quad \dots (9)$$

where  $k(x)p(\beta) = -\ln(x^2 + \beta^2)$ ,  $s(x) = -\ln(\pi)$  ,  $q(\beta) = \ln(\beta)$

since  $p(\beta) \neq 0$ , we can reduce the eq. (9) using technique of T.O.M<sup>[7\*]</sup> and we can then have the following values of parameter  $\beta$

$$\beta^i = \sqrt{\frac{x_i^2 - k_i x_{i+1}^2}{k_i - 1}}$$

where  $k_i = \frac{y_{i+1}}{y_i}$  ,  $1 \leq i \leq n-1$ .

Finally, from the previous examples we can find the estimation of parameter  $\beta$  as follows:

$$\hat{\beta} = \text{Min} \left( \sum_{m=1}^n [f(x_m, \beta^i) - y_m] \right) \quad i = 1, 2, \dots, n-1$$

## 6. RESULTS:

In Table below we show the comparison between T.O.M and MLE method of Scale parameter Cauchy distribution.

**Table:** Comparison between Scale Parameter for Cauchy Distribution between T.O.M. and MLE<sup>[8\*]</sup>.

$\theta$	N	T.O.M	MLE
0.4	10	0.3844	0.17692
	40	0.399998	0.32517
	100	0.400103	0.25687
	1000	0.398846	0.27015
1.1	10	1.059873	0.23719
	40	1.105041	0.97479
	100	1.092077	0.99639
	1000	1.096563	0.88016

---

3	10	3.002724	1.2886
	40	3.063469	2.2469
	100	3.061334	2.683
	1000	3.004707	2.76

**REFERENCES:**

- [1] Andreas V. (2010), "Notes on exponential family distributions and generalized linear models".
- [2] Watkins, Joseph C. (2009). "Exponential Families of Random Variables", University of Arizona, p: 1.
- [3] Jurgen, S. (1999). "Mathematical Statistics I", Utah State University, p: 120.
- [4] Robert V. Hogg and Allen T. Craig (1978). "Introduction to Mathematical Statistics", 4<sup>th</sup> Edition, p: 357.
- [5] Anirban DasGupta A. (2011). "Probability for Statistics and Machine Learning", Springer, pp: 583-596.
- [6] Lawrence D. Brown (1986). "Fundamentals of statistical exponential families with applications in statistical decision theory", Vol. 9, IMS Series.
- [7] Labban, J.A. (2012). "Estimation of Single Distributions Parameter by T.O.M with Exponential Families", American Journals Science, Issue 55, pp. 70-75.
- [8] Nicu S. and Michael S. Lew (2003). "Robust Computer Vision - Theory and Applications", Kluwer Academic Publishers, Volume 26, p. 54.